CSE 2194: Supervised Machine Learning
Programming Assignment-IV
(Sklearn, Keras, Linear Regression,)

Question 1:

Write a code to apply the PCA (Principal Component Analysis) algorithm on the “Digits" dataset and visualize the reduced-dimensional data using Python with sci-kit-learn and Matplotlib.
Hints: Reduce the dimensionality from 64 to 2 principal components.

Question 2:

Given a dataset containing categorical and continuous features, write a code using Python with sci-kit-learn to preprocess the data using binarization, label encoding, and label binarization to prepare it for a machine learning model. Save the final data in your local repository.

```
data = {
    "City": ["New York", "Los Angeles", "Chicago", "Miami", "Boston"],
    "Temperature (C)": [20, 25, 15, 30, 18],
    "Humidity (%)": [65, 70, 60, 75, 55],
    "Month": ["January", "February", "March", "April", "May"],
    "Target": ["Yes", "No", "Yes", "No", "Yes"]
}
```

Question 3:

Write a code to build a fully connected neural network with TensorFlow consisting of an input layer having 1000 neurons, three hidden layers with neurons H1: 512, H2: 128, H3: 64) with relu activation function, give a dropout of 15% between the second and third hidden layers, and an output layer with ten neurons and having softmax activation function. Finally, print its summary.

Question 4:

Write the explanation of exploratory data analysis.

Question 5:

Write the full calculation and answers to this question in your assignment copy.
Annual Revenue data for a company given as

Y	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
Rev. In billion Rupees	61.2	58.3	67.1	69.2	68.9	83.5	89.1	80	92.3	93	97

a) Draw a least square line fitting the data.
b) What is the expected revenue in 2025?
c) Analyze expected errors (RMSE) in predictions.

Question 6:

Write a code using the Python sci-kit-learn library for the diabetes dataset available in sci-kit-learn to implement simple linear regression. Additionally, print the weights and bias coefficients.

Question 7:

Write the Python program using sci-kit-learn to implement a housing price prediction task using multiple linear regression. The program should include steps for loading a housing dataset, preprocessing the data, training a multiple linear regression model, making predictions, and evaluating the model's performance. Data